

**IN THE CLAIMS:**

Please replace the claims with the claims provided in the listing below wherein status, amendments, additions and cancellations are indicated.

1. (Cancelled).

2. (Currently Amended) A method to predict the topology of the spatial arrangement of an amino acid sequence comprising:

using an entropy evaluation model that takes into account the global contributions of entropy to the folding of a protein (herein referred to by the name cross linking entropy (CLE) model) combined with other thermodynamic potentials as a protein-folding model to predict said topology, wherein using said entropy evaluation model to predict said topology comprises the following steps:

A. inputting an amino acid sequence of said protein,

B. preparing information on secondary structure of said amino acid sequence by way of at least one theoretical or experimental estimate,

C. applying the CLE model to said amino acid sequence and secondary structure information to evaluate the free energy of a combinatorial number of  $\beta$ -strand and  $\alpha$ -helix arrangements performed in as rapidly as

polynomial time :  $c(n-1)(n+1)$  wherein  $c$  is a constant and  $n$  is the number of secondary structure elements found in said amino acid in step A and prepared in step B,

- D. applying the CLE model in conjunction with other thermodynamic potentials that approximate hydrophobic, electrostatic and polar interactions in a thermodynamic calculation to account for both short and long range folding interactions and predict a minimum free energy and corresponding topology of the said amino acid sequence,
  - E. applying the CLE model to predict the global folding kinetics information of said amino acid sequence, and
  - F. storing the global folding kinetics information in a data file or in other form of digital memory, wherein steps A. through F. are performed with a computer.
3. (Previously Presented) A method according to claim 2, in which the cross linking entropy (CLE) model is used to evaluate the entropy loss of said protein due to folding into a particular topology given a known secondary or estimated secondary structure.
4. (Previously Presented) A method according to claim 3, in which an initial theoretical estimate of the secondary structure is obtained from either a theoretical source, or an experimental source.
5. (Previously Presented) A method according to claim 4, in which said initial theoretical estimate is from an experimental source that is an NMR experiment, an x-ray crystallography experiment, or both.

6. (Previously Presented) A method according to claim 5, in which the theoretical estimate is further supplemented with sequence alignment to find regions in which conserved segments remain essentially unchanged by differences in the aligned sequences.

7. (Currently Amended) A method to predict the topology of the spatial arrangement of an amino acid sequence, comprising:

using an entropy evaluation model that takes into account the global contributions of entropy to the folding of a protein (herein referred to by the name cross linking entropy (CLE) model) combined with other thermodynamic potentials as a protein-folding model to predict said topology, wherein

the cross linking entropy (CLE) model is used to evaluate the entropy loss of said protein due to folding into a particular topology given a known secondary or estimated secondary structure,

an initial theoretical estimate of the secondary structure is obtained from either a theoretical source, or an experimental source,

said experimental source is an NMR experiment or x-ray crystallography, or both, and

said amino acid sequence and secondary structure information is used to calculate the free energy of a combinatorial number of  $\beta$ -strand and  $\alpha$ -helix arrangements performed in as rapidly as polynomial time:  $c(n-1)(n+1)$ , wherein  $c$  is a constant and  $n$  is the number of secondary structure elements found in said amino acid;

and further comprising

storing the results of the free energy calculation in a data file or in other form of digital memory, wherein steps of the method are performed with a computer.

8. (Previously Presented) A method to predict the topology of the spatial arrangement of an amino acid sequence comprising the following steps:

- A. inputting an amino acid sequence of a protein,
  - B. preparing information on secondary structure of said amino acid sequence by way of at least one theoretical or experimental estimate,
  - C. applying a CLE model to approximate the global folding kinetics of the said amino acid sequence,
  - D. applying the CLE model to said amino acid sequence and secondary structure information to reduce the combinatorial number of  $\beta$ -strand and  $\alpha$ -helix arrangements,
  - E. applying the CLE model in conjunction with other thermodynamic potentials that approximate hydrophobic, electrostatic and polar interactions in a thermodynamic calculation to optimize the free energy to find the most thermodynamically favorable topology for said amino acid sequence,
- wherein the global free energy (FE) contribution from the CLE between two distinct amino acid residues, herein labeled  $i$  and  $j$ , is calculated by equation (1):

$$\Delta G_{ij}^{gcle} = -T \Delta S_{ij}^{gcle} = \frac{\gamma k_B T}{\xi} \left\{ \ln \left( \frac{2\gamma \xi \Delta N_{ij}}{3\lambda_{ij}^2} \right) - 1 + \frac{3\lambda_{ij}^2}{2\gamma \xi \Delta N_{ij}} \right\} \quad (1)$$

wherein  $i$  and  $j$  represent the indices of two distinct residues in said amino acid sequence, and  $j > i$ ,  $\Delta N_{ij} = j - i + 1$  expresses the number of residues separating  $i$  and  $j$ ,  $\Delta G_{ij}^{gcle}$  is the difference in the free energy contribution to the global CLE from residues  $i$  and  $j$  transitioning from the denatured (random flight) state to the native state,  $\Delta S_{ij}^{gcle}$  is the corresponding global entropy loss,  $\xi$  is the persistence length,  $\gamma$  is a dimensionless weight parameter describing the self-avoiding properties of a polymer chain,  $k_B$  is the Boltzmann constant,  $T$  is the temperature, and  $\lambda_{ij}$  (the bond gap) expresses the amino acid separation distance between the center of mass of residue  $i$  and the center of mass of residue  $j$  when both are treated as isolated molecules, and

F. storing the optimized free energy information in a data file or in other form of digital memory, wherein steps A. through F. are performed with a computer.

9. (Previously Presented) A method to predict the topology of the spatial arrangement of an amino acid sequence comprising the following steps:

- A. inputting an amino acid sequence of a protein,
- B. preparing information on secondary structure of said amino acid sequence by way of at least one theoretical or experimental estimate,
- C. applying a CLE model to approximate the global folding kinetics of the said amino acid sequence,
- D. applying the CLE model to said amino acid sequence and secondary structure information to reduce the combinatorial number of  $\beta$ -strand and  $\alpha$ -helix arrangements,
- E. applying the CLE model in conjunction with other thermodynamic potentials that approximate hydrophobic, electrostatic and polar interactions in a thermodynamic calculation to optimize the free energy to find the most thermodynamically favorable topology for said amino acid sequence,

wherein the global free energy (FE) contribution from the CLE between two distinct amino acid residues, herein labeled  $i$  and  $j$ , is calculated by equation (1):

$$\Delta G_{ij}^{gcle} = -T \Delta S_{ij}^{gcle} = \frac{\gamma k_B T}{\xi} \left\{ \ln \left( \frac{2\gamma \xi \Delta N_{ij}}{3\lambda_{ij}^2} \right) - 1 + \frac{3\lambda_{ij}^2}{2\gamma \xi \Delta N_{ij}} \right\} \quad (1)$$

wherein,  $i$  and  $j$  represent the indices of two distinct residues in said amino acid sequence, and  $j > i$ ,  $\Delta N_{ij} = j - i + 1$  expresses the number of residues separating  $i$  and  $j$ ,  $\Delta G_{ij}^{gcle}$  is the difference in the free energy contribution to the global CLE from residues  $i$  and  $j$  transitioning

from the denatured (random flight) state to the native state,  $\Delta S_{ij}^{gcle}$  is the corresponding global entropy loss,  $\xi$  is the persistence length,  $\gamma$  is a dimensionless weight parameter describing the self-avoiding properties of a polymer chain,  $k_B$  is the Boltzmann constant,  $T$  is the temperature, and  $\lambda_{ij}$  (the bond gap) expresses the amino acid separation distance between the center of mass of residue  $i$  and the center of mass of residue  $j$  when both are treated as isolated molecules, and wherein the total CLE contribution to the free energy ( $\Delta G_{cle}$ ) is calculated by equation (2):

$$\Delta G_{cle} = \Delta G_{\xi}^o + \sum_{all\_bonds(i,j)} \Delta G_{ij}^{gcle} + \sum_{i',j'} f_{i'j'}(\xi) \quad (2)$$

wherein,  $\Delta G_{ij}^{gcle}$  is defined in equation (1),  $i'$  and  $j'$  are indices specifying two secondary structure elements ( $\alpha$ -helices or  $\beta$ -strands) that are joined together by the corresponding set of bonds  $i$  and  $j$ ,  $f_{i'j'}(\xi)$  is a positive definite penalty function used to enforce topology constraints on the minimum allowed sequence length of a loop connecting two elements of secondary structure  $i'j'$  and is a function of the persistence length  $\xi$ , and  $\Delta G_{\xi}^o$  is a renormalization correction and is an integral function of  $\xi$  as shown by equation (3):

$$\Delta G_{\xi}^o = \frac{(\gamma + 1/2)Nk_B T}{D\xi} \int_1^{\xi} \left( \frac{\ln(x)}{(1-x)} + 1 \right) dx \quad (3)$$

wherein,  $N$  indicates the number of amino acids in the said sequence,  $D$  is the dimensionality of the system, the limits in the integral  $(1 \rightarrow \xi)$  indicate the change in the number of degrees of freedom from an individual amino acid residue to a cluster of  $\xi$  amino acids treated as a group (where  $\xi > 1$  amino acid and  $\xi$  need not be an integer) and  $x$  is a dummy variable in the integral substituting for  $\xi$ , and

F. storing the most thermodynamically favorable topology information in a data file or in other form of digital memory, wherein steps A through F are performed with a computer.

10. (Cancelled)

11. (Previously Presented) A method according to any one of claims 8 to 9, in which the CLE model is applied in conjunction with other derived or constructed thermodynamic potentials that approximate hydrophobic, electrostatic and polar interactions, in a thermodynamic calculation to account for both short and long range folding interactions and predict a minimum free energy and corresponding topology of said amino acid sequence.

12. - 13. (Cancelled).